

Data-driven precompensator tuning for linear parameter varying systems

Mark Butcher, Alireza Karimi *and Roland Longchamp

Automatic Control Laboratory, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland.

SUMMARY

Methods for direct data-driven tuning of the parameters of precompensators for LPV systems are developed. Since the commutativity property is not always satisfied for LPV systems, previously proposed methods for LTI systems that use this property cannot be directly adapted. When the ideal precompensator giving perfect mean tracking exists in the proposed precompensator parameterisation, the LPV transfer operators do commute and an algorithm using only two experiments on the real system is proposed. It is shown that this algorithm gives consistent estimates of the ideal parameters despite the presence of stochastic disturbances. For the more general case, when the ideal precompensator does not belong to the set of parameterised precompensators, another technique is developed. This technique requires a number of experiments equal to twice the number of precompensator parameters and it is shown that the calculated parameters minimise the mean squared tracking error. The theoretical results are demonstrated in simulation. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: Data-driven tuning, Precompensator, Linear parameter varying systems

1. Introduction

It is commonplace to use precompensators based on the inverse of the closed-loop system to improve the tracking performance of linear time-invariant (LTI) systems. This technique typically uses the inverse of a model of the closed-loop system for the precompensator. However, the model will be subject to uncertainty and when this is above a certain level the tracking performance of the system can be adversely affected [1]. In [2] a data-driven method is proposed for direct tuning of the parameters of precompensators for LTI systems. This method minimises the tracking control criterion directly using measured data, rather than passing first through a system modelling step and then minimising a criterion based on the uncertain model. This approach means that the achieved tracking is not affected by system

*Correspondence to: Automatic Control Laboratory, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland. e-mail: alireza.karimi@epfl.ch

Contract/grant sponsor: Swiss National Science Foundation; contract/grant number: 200021-116156/1

model uncertainty and leads to high tracking performance. The method is based on parameter estimation algorithms using instrumental variables, therefore making the parameter estimates asymptotically unaffected by measurement noise. Additionally, by taking advantage of the commutativity of LTI transfer operators, it only requires data from one experiment in order to tune the precompensator. In many applications, however, the LTI assumption is not satisfied e.g. certain mechatronic systems such as x-y positioning tables where the dynamics change as a function of position and consequently the method proposed in [2] cannot be applied.

A class of systems whose dynamics change as a function of the operating point are linear parameter varying (LPV) systems. For LPV systems, methods have been proposed ([3], [4]) to tune precompensators and feedforward controllers whose parameters also vary as a function of the operating point. These methods, however, are based on uncertain identified LPV models and thus, unlike direct data-driven methods, suffer from model uncertainty.

No data-driven precompensator, or feedforward controller, tuning methods for LPV systems have been proposed to the authors' knowledge. But, as is the case for LTI systems, system identification techniques for LPV systems should be adaptable to the tuning of these controllers.

Research into the problem of identifying LPV systems has been active in recent years (see e.g. [5], [6], [7], [8]). In [9] a method is proposed for the identification of the parameters of Single Input Single Output (SISO) LPV systems in input-output form. Each parameter of the system transfer operator is a linear combination of predefined, operating point dependent functions. The identification procedure is then one of identifying the coefficients multiplying these functions, which is a linear regression problem, and so can be computed using the standard least squares technique. However, as occurs in the LTI case, the least squares technique generally gives biased parameter estimates. Consistent estimates can be obtained using instrumental variables [10].

In this paper the application of instrumental variables to the problem of direct, data-driven tuning of precompensators for LPV systems is considered. It is shown that if the ideal precompensator giving zero mean tracking exists in the proposed precompensator parameterisation, the LPV transfer operators commute and a tuning technique is proposed which gives consistent estimates using measurements from just two experiments. For the more general case, where the LPV transfer operators do not commute, another algorithm is proposed requiring a number of experiments equal to twice the number of precompensator parameters. The algorithm leads to parameter estimates that converge to those that minimise the desirable mean squares criterion.

The paper is organized as follows. Notation and preliminaries are given in Section 2. The tuning scheme when the ideal precompensator exists in the precompensator parameterisation is presented in Section 3. In Section 4, the tuning method for the general case is explained. Simulation results are presented in Section 5. Finally, some concluding remarks are made in Section 6.

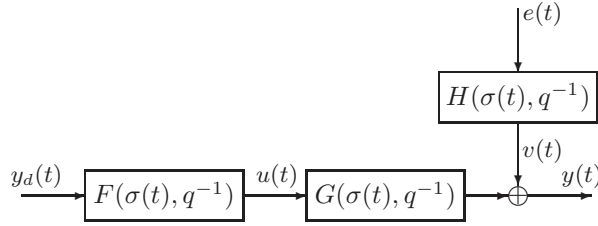


Figure 1. System with precompensator

2. Preliminaries

The output of a SISO LPV system $G(\sigma(t), q^{-1})$ is given by:

$$\begin{aligned} y(t) &= G(\sigma(t), q^{-1})u(t) + H(\sigma(t), q^{-1})e(t) \\ &= G(\sigma(t), q^{-1})u(t) + v(t) \end{aligned} \quad (1)$$

where $\sigma(t) \in \mathbb{R}^{n_\sigma}$ is a measurable scheduling parameter vector at time t , $H(\sigma(t), q^{-1})$ a, possibly LPV, transfer operator filtering the sequence of zero-mean, independent random variables $e(t)$ to give $v(t)$, and q^{-1} the backward-shift time operator. The scheduling parameter vector contains the measurable signal(s) which correspond to the system's current operating point. It should be noted that $G(\sigma(t), q^{-1})$ and $H(\sigma(t), q^{-1})$ may be the transfer operators of either an open or closed-loop system, under the condition that they are uniformly stable for all $\sigma(t)$ in the operating zone.

Definition: An LPV transfer operator $P(\sigma(t), q^{-1})$, $\sigma(t) \in \mathcal{A} \subset \mathbb{R}^{n_\sigma}$:

$$P(\sigma(t), q^{-1}) = \sum_{k=0}^{\infty} p_k(\sigma(t))q^{-k}, \quad \sigma(t) \in \mathcal{A}, \quad (2)$$

where $p_k(\sigma(t))$ is the k^{th} impulse response coefficient, is said to be *uniformly stable* if

$$|p_k(\sigma(t))| \leq p_k, \quad \text{for } t = 0, 1, \dots, \quad \sum_{k=0}^{\infty} p_k < \infty. \quad (3)$$

The output of the system with an LPV precompensator $F(\sigma(t), q^{-1})$, when the desired output $y_d(t)$ is applied at the precompensator's input, is given by (see Fig. 1):

$$y(t) = G(\sigma(t), q^{-1})F(\sigma(t), q^{-1})y_d(t) + v(t). \quad (4)$$

The objective of the paper is to calculate the parameters of the precompensator $F(\sigma(t), q^{-1})$ so as to reduce the tracking error:

$$\epsilon(t) = y_d(t) - y(t). \quad (5)$$

In this paper we consider the ideal precompensator to be that which gives zero mean tracking error. It can clearly be seen from (4) and (5) that the ideal precompensator is that for which:

$$G(\sigma(t), q^{-1})F(\sigma(t), q^{-1}) = 1. \quad (6)$$

A fact which should be noted is that due to the time-varying nature of the transfer operators, commutativity does not apply to them, in general. In fact, the backward-shift operator should obey a non-commutative multiplicative operation i.e. $q^{-i}(\sigma(t)q^{-j}u(t)) = \sigma(t-i)u(t-i-j)$.

2.1. Precompensator parameterisation

The precompensator is parameterised such that $F(\sigma(t), q^{-1})$ is linear in its parameters and can be expressed as:

$$F(\rho, \sigma(t), q^{-1}) = \beta^T(\sigma(t), q^{-1})\rho \quad (7)$$

where

$$\beta(\sigma(t), q^{-1}) = \left[\beta_0^1(q^{-1}) (\sigma(t)^T \beta_0^2(q^{-1})), \beta_1^1(q^{-1}) (\sigma(t)^T \beta_1^2(q^{-1})), \dots, \beta_{n_\beta}^1(q^{-1}) (\sigma(t)^T \beta_{n_\beta}^2(q^{-1})) \right]^T, \quad (8)$$

and $\rho \in \mathbb{R}^{n_\rho}$, with $n_\rho = (n_\sigma + 1)(n_\beta + 1)$, is the vector of controller parameters:

$$\rho = [\rho_0^0, \rho_0^1, \dots, \rho_0^{n_\sigma}, \rho_1^0, \rho_1^1, \dots, \rho_1^{n_\sigma}, \dots, \rho_{n_\beta}^0, \rho_{n_\beta}^1, \dots, \rho_{n_\beta}^{n_\sigma}]^T. \quad (9)$$

The $\beta_i^j(q^{-1})$ are linear discrete-time transfer operators, which can be any orthonormal basis functions, such as Laguerre or Kautz. In the sequel, however, for clarity of presentation, we suppose that $\beta_i^1(q^{-1}) = 1$ and $\beta_i^2(q^{-1}) = q^{-i}$. This choice means $\beta^T(\sigma(t), q^{-1})$ is given by:

$$\beta(\sigma(t), q^{-1}) = [\sigma_0(t), \sigma_1(t), \dots, \sigma_{n_\sigma}(t), \sigma_0(t)q^{-1}, \sigma_1(t)q^{-1}, \dots, \sigma_{n_\sigma}(t)q^{-1}, \dots, \sigma_0(t)q^{-n_\beta}, \sigma_1(t)q^{-n_\beta}, \dots, \sigma_{n_\sigma}(t)q^{-n_\beta}]^T \quad (10)$$

where $\sigma_j(t)$ represents the j th element of $\sigma(t)$. This parameterisation allows a wide range of dependence on the scheduling parameter to be described. For example each $\sigma_j(t)$ could represent a function of a different scheduling parameter. Alternatively the $\sigma_j(t)$ could be a set of orthogonal basis functions of a single scheduling parameter e.g. polynomials:

$$\sigma_j(t) = \bar{\sigma}^j(t), \quad (11)$$

where $\bar{\sigma}(t)$ is the single scheduling parameter.

These choices lead to the following expression for F :

$$\begin{aligned} F(\rho, \sigma(t), q^{-1}) &= [\rho_0^0 \sigma_0(t) + \rho_0^1 \sigma_1(t) + \dots + \rho_0^{n_\sigma} \sigma_{n_\sigma}(t)] \\ &\quad + [\rho_1^0 \sigma_0(t) + \rho_1^1 \sigma_1(t) + \dots + \rho_1^{n_\sigma} \sigma_{n_\sigma}(t)] q^{-1} + \dots \\ &\quad + [\rho_{n_\beta}^0 \sigma_0(t) + \rho_{n_\beta}^1 \sigma_1(t) + \dots + \rho_{n_\beta}^{n_\sigma} \sigma_{n_\sigma}(t)] q^{-n_\beta}. \end{aligned} \quad (12)$$

Remark: In the special case that the desired output $y_d(t)$ and scheduling parameter $\sigma(t)$ are known a priori, they can be used to improve the tracking of systems with a time delay. This improvement is achieved by setting $\beta_i^1(q^{-1}) = q^\delta$, where δ equals the system's time delay. This fact can be illustrated via the following example. Consider the noise-free system with a time delay m :

$$y(t) = -a_1(\sigma(t))y(t-1) - a_2(\sigma(t))y(t-2) + u(t-m). \quad (13)$$

We want $y(t) = y_d(t)$, so substituting this equality into the above equation gives:

$$u(t-m) = y_d(t) + a_1(\sigma(t))y_d(t-1) + a_2(\sigma(t))y_d(t-2) \quad (14)$$

or

$$u(t) = y_d(t+m) + a_1(\sigma(t+m))y_d(t+m-1) + a_2(\sigma(t+m))y_d(t+m-2),$$

which shows that the structure required for perfect tracking is achieved by choosing $\delta = m$. This implies that values of $\sigma(t)$ and $y_d(t)$ at $t + \delta$ should be used at time t , which is possible if they are known in advance. Unfortunately, in many applications, i.e. those where $\sigma(t)$ is measured in real time, advanced knowledge of $\sigma(t)$ will not be available.

2.2. Single realisation behaviour and ergodicity

It is often useful to be able to equate the time average properties of a signal over a single realisation with the ensemble average taken over many realisations. Signals with this property are called ergodic and Theorem 2B.1 in [11] indicates when certain types of nonstationary signals can be ergodic in the correlation.

Theorem 2B.1 [p. 55 in [11]] Let $\{P_\theta(q^{-1}), \theta \in D_\theta\}$ and $\{M_\theta(q^{-1}), \theta \in D_\theta\}$ be uniformly stable families of filters, and assume that the deterministic signal $w(t)$, $t = 1, 2, \dots$, is subject to

$$|w(t)| \leq C_w, \quad \forall t. \quad (15)$$

Let the signal $s_\theta(t)$ be defined, for each $\theta \in D_\theta$, by

$$s_\theta(t) = P_\theta(q^{-1})v_l(t) + M_\theta(q^{-1})w(t) \quad (16)$$

where

$$v_l(t) = \sum_{k=0}^{\infty} l_k(t)e_l(t-k) = L(t, q^{-1})e_l(t) \quad (17)$$

and $e_l(t)$ is a sequence of independent random vectors with zero mean values, $E\{e_l(t)e_l^T(t)\} = \Lambda_t$ and bounded fourth moments, and $\{L(t, q^{-1}), t = 1, 2, \dots\}$ is a uniformly stable family of filters. $E\{\cdot\}$ denotes the mathematical expectation. Then:

$$\sup_{\theta \in D_\theta} \left\| \frac{1}{N} \sum_{t=0}^{N-1} [s_\theta(t)s_\theta^T(t) - E\{s_\theta(t)s_\theta^T(t)\}] \right\| \rightarrow 0 \quad \text{w.p. 1, as } N \rightarrow \infty, \quad (18)$$

where $\|\cdot\|$ is the Frobenius norm.

3. Tuning when LPV transfer operators commute

As mentioned previously, in general, time-varying operators do not commute. One case, however, where they do is when the two operators considered are reciprocal. Thus, in the case that the precompensator's parameterisation and parameters are such that (6) is satisfied then $F(\rho_0, \sigma(t))G(\sigma(t)) = G(\sigma(t))F(\rho_0, \sigma(t)) = 1$, where ρ_0 are the parameters satisfying (6). This fact gives an idea for a tuning scheme for the precompensator's parameters.

3.1. Tuning scheme

We have that the tracking error of the system, with a precompensator, is given by:

$$\epsilon(t) = y_d(t) - G(\sigma(t), q^{-1})F(\rho, \sigma(t), q^{-1})y_d(t) - v(t). \quad (19)$$

In the absence of noise, and when $G(\sigma(t), q^{-1})$ and $F(\rho, \sigma(t), q^{-1})$ are commutative, the same tracking error would be obtained if the positions of the system and the precompensator were

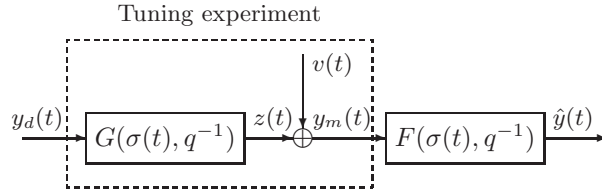


Figure 2. Precompensator tuning scheme

swapped so that F acts as a post-compensator. Using this idea, it is possible to estimate $\epsilon(t)$ from one set of data obtained from the system without a precompensator as:

$$\begin{aligned}\hat{\epsilon}(t) &= y_d(t) - \hat{y}(t) = y_d(t) - F(\rho, \sigma(t), q^{-1})y_m(t) \\ &= y_d(t) - F(\rho, \sigma(t), q^{-1})G(\sigma(t), q^{-1})y_d(t) - F(\rho, \sigma(t), q^{-1})v(t) \\ &= y_d(t) - F(\rho, \sigma(t), q^{-1})z(t) - F(\rho, \sigma(t), q^{-1})v(t)\end{aligned}\quad (20)$$

where $z(t)$ and $y_m(t)$ are the noise-free and noisy outputs, respectively, of the system when $y_d(t)$ is applied as the input (see Fig. 2).

3.2. Algorithm

It is possible to express $\hat{\epsilon}(t)$ in linear regression form as:

$$\hat{\epsilon}(t) = y_d(t) - F(\rho, \sigma(t), q^{-1})y_m(t) = y_d(t) - \phi_m^T(t)\rho \quad (21)$$

where:

$$\begin{aligned}\phi_m(t) &= [\sigma_0(t)y_m(t), \sigma_1(t)y_m(t), \dots, \sigma_{n_\sigma}(t)y_m(t), \\ &\quad \sigma_0(t)y_m(t-1), \sigma_1(t)y_m(t-1), \dots, \sigma_{n_\sigma}(t)y_m(t-1), \dots, \\ &\quad \sigma_0(t)y_m(t-n_\beta), \sigma_1(t)y_m(t-n_\beta), \dots, \sigma_{n_\sigma}(t)y_m(t-n_\beta)]^T.\end{aligned}\quad (22)$$

The precompensator parameters can then be found by the minimisation of a quadratic cost function:

$$J_m^N(\rho) = \frac{1}{2N} \sum_{t=0}^{N-1} \hat{\epsilon}^2(t), \quad (23)$$

where N is the number of data points for which $y_d(t)$ is defined. The minimiser of this criterion is given by:

$$\rho_{mLS}^N = \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_m(t)\phi_m^T(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_m(t)y_d(t). \quad (24)$$

Unfortunately when $F(\rho, \sigma(t))$ is placed as a post-compensator, it filters the noise $v(t)$ also, as seen in Fig. 2. Therefore the parameters which minimise the variance of the tracking error estimate will not be the same as those which minimise the variance of the true tracking error.

The Instrumental Variables (IV) method can be used, nonetheless, to give consistent estimates of the true minimising parameters ρ_0 . For the IV estimates to converge to the

true values, the IV vector must be correlated with the non-noisy component of $y_m(t)$, but not with the noise $v(t)$. Many choices of IV vector satisfy these conditions, such as a vector of time shifted versions of $y_d(t)$. The choice considered in this paper is to use a vector similar to $\phi_m(t)$, but with $y_m(t)$ obtained from a second experiment, performed in the same way as the first. The second experiment will, however, be affected by a different, independent noise realisation. This choice has been made as it leads to an IV vector that is strongly correlated with the non-noisy component of $y_m(t)$, and so leads to parameter estimates with low variances. The IV estimate is thus given by:

$$\rho_{mIV}^N = \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_{m2}^T(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) y_d(t), \quad (25)$$

where $\phi_{m1}(t)$ and $\phi_{m2}(t)$ are the $\phi_m(t)$ from the two experiments using $y_{m1}(t)$ and $y_{m2}(t)$ respectively.

The consistency of the IV estimate is not directly obvious as, unlike the standard LTI case, the signals considered contain nonstationary stochastic components. The applicability of ergodicity type results typically used in consistency analysis is not, therefore, immediately evident. An analysis is thus performed in the next subsection which demonstrates that the IV method does indeed lead to consistent estimates, despite the presence of these types of disturbances.

3.3. Consistency of the IV estimate

To see that the IV method gives consistent estimates we begin by rewriting (25) as:

$$\begin{aligned} \rho_{mIV}^N &= \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_{m2}^T(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_z^T(t) \rho_0 \\ &= \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_{m2}^T(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) (\phi_{m2}^T(t) - \phi_{v2}^T(t)) \rho_0 \\ &= \rho_0 - \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_{m2}^T(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_{v2}^T(t) \rho_0, \end{aligned} \quad (26)$$

where $\phi_z(t)$ and $\phi_{vi}(t)$ are similar to $\phi_{mi}(t)$, but $y_{mi}(t)$ is replaced by $z(t)$ and $v_i(t)$ respectively. Additionally $\phi_{mi}(t) = \phi_z(t) + \phi_{vi}(t)$.

In order for the parameter estimates to be consistent i.e. that ρ_{mIV}^N converges almost surely to ρ_0 as $N \rightarrow \infty$, it is necessary that:

$$\text{i) } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_{m2}^T(t) \text{ be nonsingular.} \quad (27)$$

$$\text{ii) } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_{v2}^T(t) \rho_0 = \mathbf{0}, \quad (28)$$

where $\mathbf{0}$ is the zero vector.

Condition i) is a persistency of excitation condition. It is similar to the persistency of excitation condition found in methods for the identification of input-output form LPV models,

as similar signals are involved i.e. noisy output signals multiplied by functions of the scheduling parameter. Consistency of excitation for LPV identification was first considered in [9] and sufficient conditions for polynomial type coefficient dependence on the scheduling parameter are given. It is shown that if the system input signal is ‘sufficiently rich’ then persistency of excitation is ensured if the scheduling parameter ‘visits’ $n_\sigma + 1$ distinct points infinitely many times, where n_σ is the order of the polynomial dependence. More recently [12] produced more general sufficient conditions for other types of coefficient dependence.

To show Condition ii) further analysis is required. We have that:

$$\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}(t) \phi_{v2}^T(t) = \frac{1}{N} \sum_{t=0}^{N-1} \phi_z(t) \phi_{v2}^T(t) + \frac{1}{N} \sum_{t=0}^{N-1} \phi_{v1}(t) \phi_{v2}^T(t). \quad (29)$$

Considering the first matrix on the right hand side of (29), each of its elements is the sum over time of products of terms such as $\sigma_j(t)z(t-n)$ and $\sigma_i(t)v_2(t-p)$. Then, referring to Theorem 2B.1 given in Subsection 2.2, we can define:

$$\begin{aligned} s(t) &= \begin{bmatrix} \sigma_j(t)z(t-n) \\ \sigma_i(t)v_2(t-p) \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ \sigma_i(t)H(\sigma(t-p), q^{-1})e_2(t-p) \end{bmatrix} + \begin{bmatrix} \sigma_j(t)G(\sigma(t-n), q^{-1})y_d(t-n) \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} v_1^1(t) \\ v_2^1(t) \end{bmatrix} + \begin{bmatrix} w_1^1(t) \\ w_2^1(t) \end{bmatrix} \end{aligned} \quad (30)$$

where $w_1^1(t)$, $w_2^1(t)$, $v_1^1(t)$ and $v_2^1(t)$ have their obvious definitions. The signals $w_1^1(t)$ and $w_2^1(t)$ satisfy (15), due to the assumed uniform stability of $G(\sigma(t), q^{-1})$ and the boundedness of $\sigma_j(t)$ and $y_d(t)$. Also $v_1^1(t)$ and $v_2^1(t)$ fit in with the desired form of (17) due to the assumed uniform stability of $H(\sigma(t), q^{-1})$. The components of $s(t)s^T(t)$ give, amongst others, $\sigma_j(t)z(t-n)\sigma_i(t)v_2(t-p)$. So by applying Theorem 2B.1 we can state that:

$$\left\| \frac{1}{N} \sum_{t=0}^{N-1} [\phi_z(t) \phi_{v2}^T(t) - E\{\phi_z(t) \phi_{v2}^T(t)\}] \right\| \rightarrow 0 \quad \text{w.p. 1, as } N \rightarrow \infty,$$

and, as $z(t)$ and $v(t)$ are not correlated, $E\{\phi_z(t) \phi_{v2}^T(t)\} = \mathbf{0}$ implying:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \phi_z(t) \phi_{v2}^T(t) = \mathbf{0} \quad \text{w.p. 1.}$$

A similar result for the second matrix on the right hand side of (29) can be derived using Theorem 2B.1 i.e. that:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{v1}(t) \phi_{v2}^T(t) = \mathbf{0} \quad \text{w.p. 1}$$

and combining these two satisfies Condition ii), showing the parameters estimates obtained with the IV method to be consistent.

3.4. Noisy scheduling parameters

The values of the scheduling parameter $\sigma_j(t)$ used in the calculation of ρ_{mIV}^N will often also be measured. They are, therefore, susceptible to measurement noise as well. In the case that the noise-to-signal ratio is very low, the effect of this noise can be neglected. However, if the ratio is not negligible the measurement noise may lead to biased parameter estimates.

The effect of noisy scheduling parameter measurements on the consistency of certain LPV identification methods has already been considered in [10]. It was shown that if the scheduling parameter is uncorrelated with the output signal and the dependency on it is affine, then consistent parameter estimates are obtained using an IV vector with scheduling parameter values measured during a second experiment. It will be shown in this section that a similar result holds for the consistency of the estimated precompensator parameters.

We express the noisy, measured scheduling parameter vector $\sigma_v(t)$ as the sum of a noise-free component and a noisy component i.e.

$$\sigma_v(t) = \sigma(t) + v_\sigma(t). \quad (31)$$

$\sigma_v(t) \in \mathbb{R}^{n_\sigma}$ and $v_\sigma(t) \in \mathbb{R}^{n_\sigma}$ are, therefore, stochastic vectors, and additionally $v_\sigma(t)$ is assumed zero-mean and to have independent realisations. The measured regressor vector is now given by:

$$\phi_{mi}^v(t) = \phi_{mi}(t) + \phi_{mi}^\sigma(t), \quad (32)$$

where $\phi_{mi}(t)$ is as defined in (22) for experiment i and

$$\phi_{mi}^\sigma(t) = [v_{\sigma i}^T(t)y_{mi}(t), v_{\sigma i}^T(t)y_{mi}(t-1), \dots, v_{\sigma i}^T(t)y_{mi}(t-n_\beta)]^T. \quad (33)$$

The IV estimate is then given by:

$$\rho_{mIV}^N = \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t) \phi_{m2}^{v^T}(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t) y_d(t). \quad (34)$$

3.4.1. Consistency of the IV estimate with noisy scheduling parameters

To show that using this estimate consistent estimates can be obtained we start by rewriting (34) as:

$$\begin{aligned} \rho_{mIV}^N &= \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t) \phi_{m2}^{v^T}(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t) \phi_z^T(t) \rho_0 \\ &= \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t) \phi_{m2}^{v^T}(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t) (\phi_{m2}^v(t) - \phi_{m2}^\sigma(t) - \phi_{v2}(t))^T \rho_0 \\ &= \rho_0 - \left[\frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t) \phi_{m2}^{v^T}(t) \right]^{-1} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t) (\phi_{m2}^\sigma(t) - \phi_{v2}(t))^T \rho_0. \end{aligned} \quad (35)$$

For the estimate to be consistent we require, in a similar fashion to before, that:

$$\text{i) } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \phi_{m1}^v(t) \phi_{m2}^{v^T}(t) \text{ be nonsingular.} \quad (36)$$

$$\text{ii) } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \phi_{m1}^v(t) (\phi_{m2}^\sigma(t) - \phi_{v2}(t))^T = 0. \quad (37)$$

As before, Condition i) is a persistency of excitation condition.

Condition ii) can, again, be analysed using Theorem 2B.1. We first consider the matrix $\phi_{m1}^v(t)\phi_{m2}^{\sigma^T}(t)$. This contains the following cross-product type terms:

$$\sigma_{v1,j}(t)y_{m1}(t-p)v_{\sigma2,i}(t)y_{m2}(t-n) = [\sigma_j(t) + v_{\sigma1,j}(t)]y_{m1}(t-p)v_{\sigma2,i}(t)y_{m2}(t-n). \quad (38)$$

Considering (38), it is clearly equal to:

$$\begin{aligned} & \sigma_j(t) [G(\sigma(t-p), q^{-1})y_d(t-p) + H(\sigma(t-p), q^{-1})e_1(t-p)] \\ & v_{\sigma2,i}(t) [G(\sigma(t-n), q^{-1})y_d(t-n) + H(\sigma(t-n), q^{-1})e_2(t-n)] \\ & + v_{\sigma1,j}(t) [G(\sigma(t-p), q^{-1})y_d(t-p) + H(\sigma(t-p), q^{-1})e_1(t-p)] \\ & v_{\sigma2,i}(t) [G(\sigma(t-n), q^{-1})y_d(t-n) + H(\sigma(t-n), q^{-1})e_2(t-n)]. \end{aligned} \quad (39)$$

So, referring to Theorem 2B.1, we can write:

$$\begin{aligned} s_2(t) &= \begin{bmatrix} 0 \\ \sigma_j(t)H(\sigma(t-p), q^{-1})e_1(t-p) \\ v_{\sigma2,i}(t)G(\sigma(t-n), q^{-1})y_d(t-n) \\ v_{\sigma2,i}(t)H(\sigma(t-n), q^{-1})e_2(t-n) \\ v_{\sigma1,j}(t)G(\sigma(t-p), q^{-1})y_d(t-p) \\ v_{\sigma1,j}(t)H(\sigma(t-p), q^{-1})e_1(t-p) \end{bmatrix} + \begin{bmatrix} \sigma_j(t)G(\sigma(t-p), q^{-1})y_d(t-p) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} v_1^2(t) \\ v_2^2(t) \\ v_3^2(t) \\ v_4^2(t) \\ v_5^2(t) \\ v_6^2(t) \end{bmatrix} + \begin{bmatrix} w_1^2(t) \\ w_2^2(t) \\ w_3^2(t) \\ w_4^2(t) \\ w_5^2(t) \\ w_6^2(t) \end{bmatrix}. \end{aligned} \quad (40)$$

In order to write $v_4^2(t) = v_{\sigma2,i}(t)H(\sigma(t-n), q^{-1})e_2(t-n)$ and $v_6^2(t) = v_{\sigma1,j}(t)H(\sigma(t-p), q^{-1})e_1(t-p)$ it is necessary that $v_{\sigma,j}(t)$ and $e(t)$ are uncorrelated. This condition is reasonable so long as the scheduling parameter is not the system output $y(t)$. If they are correlated the expected value of their product is non-zero and does not satisfy the theorem's assumptions on the stochastic component. The ergodicity of the signals used in the identification method is, thus, not provable in this case.

Additionally it is not possible to establish the ergodicity of the signals when the scheduling parameter has the polynomial dependence discussed in (11) which is of an order greater than 1. The reason is that we would have higher order moments of the noise term affecting $\bar{\sigma}(t)$, which are non-zero mean. This, in turn, would imply that $v_{\sigma,j}(t)$ is non-zero mean and thus $v_4^2(t)$ and $v_6^2(t)$ are non-zero mean, violating the theorem's assumptions.

With these conditions in mind, we see that amongst the elements of $s_2(t)s_2^T(t)$ are all the cross-terms found in (39), and thus:

$$\left\| \frac{1}{N} \sum_{t=1}^N [\sigma_{v1,j}(t)y_{m1}(t-p)v_{\sigma2,i}(t)y_{m2}(t-n) - E\{\sigma_{v1,j}(t)y_{m1}(t-p)v_{\sigma2,i}(t)y_{m2}(t-n)\}] \right\| \rightarrow 0 \quad \text{w.p. 1, as } N \rightarrow \infty. \quad (41)$$

Now since the vector $v_\sigma(t)$ is assumed zero-mean, uncorrelated with the system output and to have independent realisations we have that:

$$E\{\sigma_{v1,j}(t)y_{m1}(t-p)v_{\sigma2,i}(t)y_{m2}(t-n)\} = E\{\sigma_{v1,j}(t)y_{m1}(t-p)y_{m2}(t-n)\}E\{v_{\sigma2,i}(t)\} = 0. \quad (42)$$

This result implies that:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \sigma_{v1,j}(t)y_{m1}(t-p)v_{\sigma2,i}(t)y_{m2}(t-n) = 0 \text{ w.p. } 1 \quad (43)$$

which in turn implies:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t)\phi_{m2}^{\sigma T}(t) = \mathbf{0} \text{ w.p. } 1. \quad (44)$$

We next consider the second matrix in (37) $\phi_{m1}^v(t)\phi_{v2}^T(t)$. This contains the following cross-product type terms:

$$\sigma_{v1,j}(t)y_{m1}(t-p)\sigma_i(t)v_2(t-n). \quad (45)$$

Under similar conditions on the scheduling parameter, Theorem 2B.1 can be used to show that:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \sigma_{v1,j}(t)y_{m1}(t-p)\sigma_i(t)v_2(t-n) = 0 \text{ w.p. } 1 \quad (46)$$

and so:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} \phi_{m1}^v(t)\phi_{v2}^T(t) = \mathbf{0} \text{ w.p. } 1. \quad (47)$$

(44) and (47) imply that Condition ii) is satisfied by the IV estimate (34) so when Condition i) is also satisfied the estimate is consistent despite the contamination of the scheduling parameter measurements by noise.

4. Tuning when LPV transfer operators do not commute

In general condition (6) is unlikely to be satisfied by a precompensator with the linear parameterisation proposed. In this case, the order of the precompensator can be increased until the condition is approximately satisfied and the method of the previous section can be used. In practice, however, the order required to approximately satisfy (6) may be too large to be implemented on the real system. A method that can be used in this situation is therefore developed below.

The signal $u(t) = F(\rho, \sigma(t))y_d(t)$ can be expressed as:

$$u(t) = F(\rho, \sigma(t))y_d(t) = \phi^T(t)\rho \quad (48)$$

where

$$\begin{aligned} \phi^T(t) = & [\sigma_0(t)y_d(t), \sigma_1(t)y_d(t), \dots, \sigma_{n_\sigma}(t)y_d(t), \\ & \sigma_0(t)y_d(t-1), \sigma_1(t)y_d(t-1), \dots, \sigma_{n_\sigma}(t)y_d(t-1), \dots, \\ & \sigma_0(t)y_d(t-n_\beta), \sigma_1(t)y_d(t-n_\beta), \dots, \sigma_{n_\sigma}(t)y_d(t-n_\beta)]. \end{aligned} \quad (49)$$

It is now possible to write the output as:

$$\begin{aligned} y(t) &= G(\sigma(t), q^{-1})F(\rho, \sigma(t), q^{-1})y_d(t) + v(t) \\ &= G(\sigma(t), q^{-1})\phi^T(t)\rho + v(t) \\ &= x^T(t)\rho + v(t), \end{aligned} \quad (50)$$

where $x(t) = G(\sigma(t), q^{-1})\phi(t) \in \mathbb{R}^{n_\rho}$.

The tracking error (5) can then be expressed as:

$$\epsilon(t) = y_d(t) - x^T(t)\rho - v(t) \quad (51)$$

or in vector form as:

$$\epsilon = y_d - X\rho - v, \quad (52)$$

where the vector ϵ is given by:

$$\epsilon = [\epsilon(0), \epsilon(1), \dots, \epsilon(N-1)]^T, \quad (53)$$

y_d and v are defined similarly, and the matrix X is:

$$X = [x(0), x(1), \dots, x(N-1)]^T. \quad (54)$$

The aim is to find ρ such that the average tracking error is small, therefore a logical objective is to minimise its mean squared value i.e. to find the ρ that minimises:

$$J^N(\rho) = \frac{1}{2N} E \{ \epsilon(\rho)^T \epsilon(\rho) \}, \quad (55)$$

where the dependence of ϵ on ρ has been shown explicitly.

The minimiser of (55) is given by:

$$\rho_{LMS}^N = \left[\frac{1}{N} X^T X \right]^{-1} \frac{1}{N} X^T y_d. \quad (56)$$

In the case that $G(\sigma(t), q^{-1})$ were known exactly, X could be calculated, followed by ρ_{LMS}^N . $G(\sigma(t), q^{-1})$ is never known exactly, however, and the model uncertainty leads to a non optimal ρ .

It is possible to obtain an estimate \hat{X} of the matrix X without the use of a model through a series of experiments on the real system. This can be seen to be the case by noting that the element $X_{t,j}$ of the matrix X is the output of the system $G(\sigma(t), q^{-1})$ when the j th element of $\phi(t)$ is applied as an input. Thus for each column of X an experiment can be carried out on the real system i.e. n_ρ experiments in total. In reality an estimate, rather than the exact value, of X will be found as each experiment will have its own noise realisation $v_j(t)$. The estimate of X is:

$$\hat{X} = X + V, \quad (57)$$

where V is a matrix whose t, j th element is $v_j(t)$. Substituting \hat{X} in for X in (56) we have:

$$\begin{aligned} \hat{\rho}_{LMS}^N &= \left[\frac{1}{N} (X + V)^T (X + V) \right]^{-1} \frac{1}{N} (X + V)^T y_d \\ &= \left[\frac{1}{N} (X^T X + V^T X + X^T V + V^T V) \right]^{-1} \frac{1}{N} (X + V)^T y_d. \end{aligned} \quad (58)$$

Therefore when \hat{X} is used in place of X in (56) the presence of the noise in the experiments performed to find \hat{X} will mean that the minimising value ρ_{LMS}^N cannot be calculated i.e. $\hat{\rho}_{LMS}^N \neq \rho_{LMS}^N$.

One way of dealing with this problem is to use instrumental variables again. This time two estimates of X , \hat{X}_1 and \hat{X}_2 , are used. They are obtained from two sets of the n_p experiments previously described. The IV estimate is then given by:

$$\hat{\rho}_{IV}^N = \left[\frac{1}{N} \hat{X}_1^T \hat{X}_2 \right]^{-1} \frac{1}{N} \hat{X}_1^T y_d. \quad (59)$$

Remarks:

1. The same $G(\sigma(t), q^{-1})$ should be used at each t for all the experiments i.e. when different signals $u(t)$ are applied. It is therefore necessary that the scheduling parameter $\sigma(t)$ be independent of $u(t)$, which may not be the case for certain quasi-LPV systems where the scheduling parameter can be an input-dependent, internal variable. Nonetheless, a number of LPV systems found in practice do satisfy this requirement, such as x-y positioning tables where the dynamics of one stage depend on the position of the other and not their own.
2. The specific choice of instrumental variables obtained from a second set of experiments is used because it can be shown that consistent estimates of the true minimiser of (55) are obtained using them, as will be shown in the next subsection. This is not necessarily possible for a general choice of instrumental variables which are just correlated with the system input and not with noise affecting the output. A disadvantage of this specific choice is that, due to the second set of noise realisations affecting the instrumental variables, the variance of the estimates can be large when the number of data measured is small.

4.1. Consistency of estimates

To demonstrate that the IV estimate (59) converges to the true minimiser of (55) when $N \rightarrow \infty$ we first write:

$$\begin{aligned} \hat{\rho}_{IV}^N &= \left[\frac{1}{N} (X + V_1)^T (X + V_2) \right]^{-1} \frac{1}{N} (X + V_1)^T y_d \\ &= \left[\frac{1}{N} (X^T X + V_1^T X + X^T V_2 + V_1^T V_2) \right]^{-1} \frac{1}{N} (X + V_1)^T y_d \end{aligned} \quad (60)$$

where V_1 and V_2 denote matrices of noise realisations associated with the first and second set of experiments. For $\hat{\rho}_{IV}^N$ to converge to ρ_{LMS}^N we require $V_1^T X + X^T V_2 + V_1^T V_2$ and $V_1^T y_d$ to converge to zero as $N \rightarrow \infty$. This can be shown to be the case by noting that the i, j th element of the matrix $V_1^T X$ is given by:

$$\begin{aligned} [V_1^T X]_{i,j} &= \sum_{t=0}^{N-1} v_{1i}(t) X_{t,j} \\ &= \sum_{t=0}^{N-1} [H(\sigma(t), q^{-1}) e_{1i}(t)] [G(\sigma(t), q^{-1}) \phi_j(t)], \end{aligned} \quad (61)$$

where $v1_i(t)$ corresponds to the t, i element of V_1 , and $v1_i(t) = H(\sigma(t), q^{-1})e1_i(t)$. Theorem 2B.1 is applicable to this time sum and can be used to show that:

$$\left\| \frac{1}{N} \sum_{t=0}^{N-1} [v1_i(t)X_{t,j} - E\{v1_i(t)X_{t,j}\}] \right\| \rightarrow 0 \quad \text{w.p. 1, as } N \rightarrow \infty, \quad (62)$$

and since $E\{v1_{t,i}X_{t,j}\} = 0$ we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} v1_{t,i}X_{t,j} = 0 \quad \text{w.p. 1}$$

This result can be applied to each element of $V_1^T X$, and similar results hold for $X^T V_2$, $V_1^T V_2$ and $V_1^T y_d$. Therefore, $\hat{\rho}_{IV}^N \rightarrow \rho_{LMS}^N$ as $N \rightarrow \infty$ almost surely.

4.2. Noisy scheduling parameters

As was the case for the method in the previous section, it is possible that the values of scheduling parameters used will have been measured and thus contaminated by noise. For example consider the case when the precompensator is tuned to control the lower axis of an x-y positioning table whose dynamics depend on the position of the upper axis. In this case the scheduling parameter would be the position of the upper axis. This can either be the position reference signal or the measured position. The former may not be a very good measure of the true position and the latter is affected by noise. When the noise is negligible its effect can be neglected and the IV estimate (59) can be used. However, when this is not the case, it will be shown that, under certain conditions on the scheduling parameter, by using two independent sets of scheduling parameter measurements, the IV estimate will lead to consistent estimates.

We have that the noisy, measured scheduling parameter vector $\sigma_v(t)$ is expressed as the sum of a noise-free component and a noisy component as in (31). When this is used to construct the $\phi(t)$ vector, the noisy $\phi(t)$ for the i^{th} set of experiments is given by:

$$\phi_i^v(t) = \phi_i(t) + \phi_i^\sigma(t), \quad (63)$$

where $\phi_i(t)$ is as defined in (49) for the set of experiments i and

$$\phi_i^\sigma(t) = [v_{\sigma i}^T(t)y_d(t), v_{\sigma i}^T(t)y_d(t-1), \dots, v_{\sigma i}^T(t)y_d(t-n_\beta)]^T. \quad (64)$$

It is therefore the elements of $\phi_i^v(t)$ that are now used as the inputs for the n_ρ experiments. This will lead to the new estimate of X :

$$\hat{X}_i^v = X + X_i^\sigma + V_i. \quad (65)$$

The IV estimate is now:

$$\hat{\rho}_{IV}^N = \left[\frac{1}{N} \hat{X}_1^{v^T} \hat{X}_2^v \right]^{-1} \frac{1}{N} \hat{X}_1^{v^T} y_d. \quad (66)$$

4.2.1. Consistency of the IV estimate with noisy scheduling parameters

To demonstrate that the IV estimate (66) converges to the true minimiser of (55) when

$N \rightarrow \infty$ we first write:

$$\begin{aligned}\hat{\rho}_{IV}^N &= \left[\frac{1}{N} (X + X_1^\sigma + V_1)^T (X + X_2^\sigma + V_2) \right]^{-1} \frac{1}{N} (X + X_1^\sigma + V_1)^T y_d \\ &= \left[\frac{1}{N} (X^T X + X^T X_2^\sigma + X^T V_2 + X_1^\sigma X + X_1^\sigma X_2^\sigma + X_1^\sigma V_2 + V_1^T X + V_1^T X_2^\sigma + V_1^T V_2) \right]^{-1} \\ &\quad \times \frac{1}{N} (X + X_1^\sigma + V_1)^T y_d\end{aligned}\quad (67)$$

For $\hat{\rho}_{IV}^N$ to converge to ρ_{LMS}^N we require $X^T X_2^\sigma + X^T V_2 + X_1^\sigma X + X_1^\sigma X_2^\sigma + X_1^\sigma V_2 + V_1^T X + V_1^T X_2^\sigma + V_1^T V_2$ and $(X_1^\sigma + V_1)^T y_d$ to converge to zero as $N \rightarrow \infty$. In a similar analysis to that previously presented in Subsection 4.1 this can be shown to be the case under the assumptions that the scheduling parameter is not correlated with the output and the its dependence is affine.

4.3. Input weighting

In practice the precompensator that minimises the tracking error may lead to the saturation of certain signals in the real system, most notably the plant input. It can, therefore, be of interest to be able to influence the magnitude of this input signal when tuning the parameters of a precompensator. In general, the input that is at risk of saturating, and therefore in need of reducing, can be described as:

$$\begin{aligned}u_W(t) &= U(\sigma(t), q^{-1})u(t) + H_u(\sigma(t), q^{-1})e(t) \\ &= U(\sigma(t), q^{-1})u(t) + v_u(t).\end{aligned}\quad (68)$$

It should be noted that in the case where $G(\sigma(t), q^{-1})$ describes an open-loop, stable system, it is normally the signal $u(t)$ that we wish to influence and so $U(\sigma(t), q^{-1}) = 1$ and $H_u(\sigma(t), q^{-1}) = 0$. In the case that $G(\sigma(t), q^{-1})$ represents a closed-loop system $u_W(t)$ may be an internal input, such as the plant input, so $U(\sigma(t), q^{-1})$ will have another form and $H_u(\sigma(t), q^{-1})$ will be non-zero.

It is possible to affect the magnitude of $u_W(t)$ by including a term weighting on it in the cost function:

$$J^N(\rho) = \frac{1}{2N} [E \{ \epsilon(\rho)^T W_\epsilon \epsilon(\rho) \} + E \{ u_W^T(\rho) W_u u_W(\rho) \}], \quad (69)$$

where W_ϵ and W_u are positive- and positive-semidefinite weighting matrices, respectively, and u_W is defined similarly to ϵ in (53).

Additionally, similarly to (50), we can write:

$$\begin{aligned}u_W(t) &= U(\sigma(t), q^{-1})\phi^T(t)\rho + v_u(t) \\ &= x_u^T(t)\rho + v_u(t).\end{aligned}\quad (70)$$

Therefore we can write:

$$u_W(\rho) = \begin{bmatrix} x_u^T(0) \\ x_u^T(1) \\ \vdots \\ x_u^T(N-1) \end{bmatrix} \rho = X_u \rho.$$

The minimiser of (69) is thus:

$$\rho_{LMS}^N = \left[\frac{1}{N} (X^T W_\epsilon X + X_u^T W_u X_u) \right]^{-1} \frac{1}{N} X^T y_d. \quad (71)$$

In the same way that X is not known exactly, neither is X_u , but an estimate \hat{X}_u can be obtained using the same experiments used to estimate X . This time not only the system output $y(t)$ should be measured during the experiments but also the input to be weighted $u_W(t)$. Because this input can also be affected by noise, it is necessary to use measurements from two experiments again and use the IV estimate:

$$\hat{\rho}_{IV}^N = \left[\frac{1}{N} (\hat{X}_1^T W_\epsilon \hat{X}_2 + \hat{X}_{u1}^T W_u \hat{X}_{u2}) \right]^{-1} \frac{1}{N} \hat{X}_1^T y_d. \quad (72)$$

It should be noted that input weighting was not considered for the method in Section 3, as there the aim was to identify the precompensator that satisfies condition (6) exactly. This would not be possible if a trade-off is made with the reduction of the input signal amplitude.

5. Simulation results

Simulations are used to demonstrate the effectiveness of the proposed methods.

5.1. Simulation 1

The first method, presented in Section 3, is applied to a system represented by the following transfer operator:

$$G(\sigma(t), q^{-1}) = \frac{1}{a_0(\sigma(t)) + a_1(\sigma(t))q^{-1} + a_2(\sigma(t))q^{-2}} \quad (73)$$

where

$$\begin{aligned} a_0(\sigma(t)) &= 1 - 0.2\bar{\sigma}(t) + 0.7\bar{\sigma}(t)^2, \\ a_1(\sigma(t)) &= 1 - 0.5\bar{\sigma}(t) + 0.2\bar{\sigma}(t)^2, \\ a_2(\sigma(t)) &= 1 - 0.7\bar{\sigma}(t) - 0.1\bar{\sigma}(t)^2. \end{aligned}$$

The precompensator satisfying condition (6) and giving perfect mean tracking can be represented exactly using the parameterisation proposed and thus the first method can be employed to find the correct parameters; they being:

$$\rho_0 = [1, -0.2, 0.7, 1, -0.5, 0.2, 1, -0.7, -0.1]^T. \quad (74)$$

The desired output $y_d(t)$ is a pseudo random binary signal (PRBS) with an amplitude of 1 and a shift register length of 12, giving $N = 4095$. The noise transfer operator is given by $H(\sigma(t), q^{-1}) = 1 - G(\sigma(t), q^{-1})$, as could be the case when $G(\sigma(t), q^{-1})$ represents the complementary sensitivity function of a closed-loop system. Its input $e(t)$ is taken as a zero-mean, normally distributed, stationary white noise with a variance of 0.015. The scheduling parameter is chosen as $\bar{\sigma}(t) = 0.5 + 0.4 \sin\left(\frac{2\pi t}{3}\right)$. This choice of $\bar{\sigma}(t)$ satisfies the persistency

Table I. Results of Simulation 1

	$\hat{E}\{(\rho_0 - \rho^N)^T\}\hat{E}\{\rho_0 - \rho^N\}$	$\hat{E}\{(\rho_0 - \rho^N)^T(\rho_0 - \rho^N)\}$	$\sum_{i=0}^{N-1} \hat{E}\{\epsilon_i^2\}$
mLS	0.03471	0.0457	4.1872
mIV	0.0000416	0.0123	0.3675

of excitation conditions in [9] as it visits $(n_\sigma + 1)$ distinct points periodically. The $v(t)$ resulting from these choices has an average standard deviation of 0.1287, where the average standard deviation is calculated as $(\frac{1}{N} \sum_{t=0}^{N-1} \hat{E}\{v^2(t)\})^{1/2}$. The estimated expected value \hat{E} is evaluated using the ensemble average over 200 simulations, each with a new, independent noise realisation.

Both the least squares and IV variants of the first method are tested. The orders corresponding to the precompensator satisfying (6) are used i.e. $n_\beta = 2$ and $n_\sigma = 2$. Measures of the mean and variance of the distance between the true and estimated parameters are presented in Table I. Additionally the mean 2-norm of the tracking error achieved in the absence of noise using the calculated precompensators is shown.

It can be seen that, due to its reduced sensitivity to the presence of stochastic disturbances, the IV method gives better estimates of the true parameters and thus produces precompensators that give better tracking performance.

5.2. Simulation 2

A second simulation is carried out to test the effect of noisy scheduling parameter measurements on the precompensator parameters estimated using the first method, presented in Section 3.

The system used in this simulation is given by:

$$G(\sigma(t), q^{-1}) = \frac{1}{a_0(\sigma(t)) + a_1(\sigma(t))q^{-1}} \quad (75)$$

where

$$a_0(\sigma(t)) = 1 - 0.2\bar{\sigma}(t) \quad \text{and} \quad a_1(\sigma(t)) = 1 - 0.5\bar{\sigma}(t).$$

This system is chosen because its coefficients depend affinely on the scheduling parameter. As discussed in Subsection 3.4, it is not possible to obtain consistent estimates when the scheduling parameter measurements are noisy if the dependency is not affine. The desired output $y_d(t)$ is the same signal as that used in Simulation 1. The true scheduling parameter is given by $\bar{\sigma}(t) = 0.5 + 0.4\cos(\pi t)$, which satisfies the persistency of excitation conditions in [9]. The noise $v_\sigma(t)$ contaminating the measurements of the scheduling parameter signal is a zero-mean, normally distributed, stationary white noise with a variance of 0.01. The disturbance $v(t)$ affecting the output is set to zero in order to better emphasise the effect of the scheduling parameter noise. Both the least squares and IV variants of the first method are tested. This time the IV vector uses $\sigma_j(t)$ measured during the second experiment. The values $n_\beta = 1$ and $n_\sigma = 1$ are used as they correspond to the precompensator giving perfect mean tracking whose

Table II. Results of Simulation 2

	$\hat{E}\{(\rho_0 - \rho^N)^T\} \hat{E}\{\rho_0 - \rho^N\}$	$\hat{E}\{(\rho_0 - \rho^N)^T (\rho_0 - \rho^N)\}$	$\sum_{i=0}^{N-1} \hat{E}\{\epsilon_i^2\}$
mLS	2.0588×10^{-3}	2.0776×10^{-3}	4.6624
mIV	2.8789×10^{-8}	2.4059×10^{-5}	0.0825

parameters are:

$$\rho_0 = [1, -0.2, 1, -0.5]^T. \quad (76)$$

The results are shown in Table II. Again, the expected value is estimated using the ensemble average over 200 simulations.

It is clearly evident that, again, the IV variant gives superior parameter estimates over the LS method. In turn this leads to much better tracking performance.

5.3. Simulation 3

A third simulation is performed to see how the methods behave when condition (6) cannot be satisfied using the parameterisation proposed. The system used in this simulation is given by:

$$G(\sigma(t), q^{-1}) = \frac{b_0(\sigma(t)) + b_1(\sigma(t))q^{-1}}{a_0(\sigma(t)) + a_1(\sigma(t))q^{-1}} \quad (77)$$

where

$$\begin{aligned} b_0(\sigma(t)) &= 0.5 - 0.4\bar{\sigma}(t), & b_1(\sigma(t)) &= 0.2 - 0.3\bar{\sigma}(t) \\ a_0(\sigma(t)) &= 1 - 0.2\bar{\sigma}(t), & a_1(\sigma(t)) &= 1 - 0.5\bar{\sigma}(t). \end{aligned} \quad (78)$$

The desired output $y_d(t)$ and scheduling parameter $\bar{\sigma}(t)$ are the same signals as those used in Simulation 2. The values $n_\beta = 1$ and $n_\sigma = 1$ are chosen.

A simulation is first carried out without any disturbances in order to see how the methods perform. As it is not possible to satisfy condition (6) with the finite order precompensator used, the first method is not theoretically applicable. It is tested, nonetheless, to see how it performs. The results are shown in Table III. As the parameters corresponding to the ideal precompensator do not exist in this case, the distance between those minimising the LMS criterion, calculated using (56), and those estimated is shown instead. The IV methods are not tested because in the absence of stochastic disturbances they give the same values.

From Table III we can observe that, as expected, the precompensator calculated with the general method, presented in Section 4, gives a smaller tracking error. Furthermore the first method does not find the parameters that minimise the LMS criterion and actually gives parameters that produce a larger tracking error than that obtained without a precompensator, clearly illustrating the need for the general method.

The second method is then tested in the presence of disturbances. The noise transfer operator is, again, given by $H(\sigma(t), q^{-1}) = 1 - G(\sigma(t), q^{-1})$ and $e(t)$ is a zero-mean, normally distributed white noise with a variance of 0.015. The average standard deviation of $v(t)$ this time is 0.1063. The simulations are again carried out 200 times. The results are shown in Table IV.

Table III. Results of Simulation 3 in the absence of disturbances

	$(\rho_{LMS}^N - \rho^N)^T (\rho_{LMS}^N - \rho^N)$	$\sum_{i=0}^{N-1} \epsilon_i^2$
Without precomp.	-	3082.5
mLS	10.8023	7605.0
LMS	0	1331.0

Table IV. Results of Simulation 3 in the presence of disturbances

	$\hat{E}\{(\rho_{LMS}^N - \rho^N)^T\} \hat{E}\{\rho_{LMS}^N - \rho^N\}$	$\hat{E}\{(\rho_{LMS}^N - \rho^N)^T (\rho_{LMS}^N - \rho^N)\}$	$\sum_{i=0}^{N-1} \hat{E}\{\epsilon_i^2\}$
LMS	35.0489	35.0564	2629.5
IV	0.1912	7.9891	1639.7

We see that the presence of the stochastic disturbances causes a large bias in the parameter estimates found using the LMS technique. This bias, in turn, greatly deteriorates the tracking performance obtained. The IV technique, however, is much less sensitive to the presence of the disturbances and the average values of the parameters calculated are much closer to the true minimising parameters. The negative influence of the disturbances on the variance of the parameter estimates can be seen to be greater for the IV technique than the LMS technique due to the second set of noise realisations. Nonetheless, the IV technique still gives better estimates and an average tracking error that is much closer to the value found in the absence of disturbances.

6. Conclusions

Direct data-driven tuning methods of precompensators for LPV systems are developed in this paper. These methods are model-free and minimise the tracking control criterion directly using measured data. They are, therefore, not affected by model uncertainty. This fact allows high tracking performance to be achieved compared to model-based methods. The measured data is, however, affected by stochastic disturbances which will, in general, cause biased estimates of the precompensator parameters. Instrumental variables are, therefore, used to produce consistent estimates.

Two techniques are proposed. The first one, applicable when the precompensator and the system commute, only requires two experiments in order to obtain consistent estimates of the parameters that give perfect mean tracking. The second one, which does not require that the precompensator and the system commute, uses a number of experiments equal to twice the number of precompensator parameters. It is demonstrated that the computed parameters converge to those minimising the mean squared tracking error in the presence of noise.

The main benefit of the first method over the second method is clearly the reduced number of experiments required. However, it is only applicable to the reduced class of LPV systems that commute with the precompensator. The second method is more general as does not have this condition and thus is more practically applicable. The use of sufficiently long data lengths is, however, important in order to keep the variance of the parameter estimates small.

REFERENCES

1. Devasia S. Should model-based inverse inputs be used as feedforward under plant uncertainty? *IEEE Transactions on Automatic Control*, **47**(11):1865–1871, 2002.
2. Karimi A, Butcher M, Longchamp R. Model-free precompensator tuning based on the correlation approach. *IEEE Transactions on Control Systems Technology*, **16**(5):1013–1020, September 2008.
3. Ferreres G, Roos C. Efficient convex design of robust feedforward controllers. In *Proceedings of the 44th IEEE Conference on Decision and Control*, 6460–6465, Seville, Spain, December 2005.
4. Prempain E, Postlethwaite I. Feedforward control: a full information approach. *Automatica*, **37**:17–28, 2001.
5. Lee LH, Poola K. Identification of linear parameter-varying systems using nonlinear programming. *Journal of Dynamic Systems, Measurement and Control*, **121**:71–78, 1999.
6. Verdult V, Verhaegen M. Subspace identification of multivariable linear parameter-varying systems. *Automatica*, **38**:805–814, 2002.
7. Szaier M, Mazzaro MC. An LMI approach to control-orientated identification and model (in)validation of LPV systems. *IEEE Transactions on Automatic Control*, **48**(9):1619–1624, September 2003.
8. Tóth R, Heuberger PSC, Van den Hof PMJ. LPV system identification with globally fixed orthonormal basis functions. In *Proceedings of the 46th IEEE Conference on Decision and Control*, 3646–3653, New Orleans, USA, December 2007.
9. Bamieh B, Giarré B. Identification of linear parameter varying models. *International Journal of Robust and Nonlinear Control*, **12**:841–853, 2002.
10. Butcher M, Karimi A, Longchamp R. On the consistency of certain identification methods for linear parameter varying systems. In *Proceedings of the 17th IFAC World Congress*, Seoul, Korea, July 2008.
11. Ljung L. *System Identification - Theory for the User*. Prentice Hall, NJ, USA, second edition, 1999.
12. Wei X, Del Re L. On persistent excitation for parameter estimation of quasi-LPV systems and its application in modeling of diesel engine torque. In *Proceedings of the 14th IFAC Symposium on System Identification*, 517–522, Newcastle, Australia, March 2006.